PNAS
www.pnas.org

Supplementary Information for

Diversity begets diversity: chemical and microbial diversity co-vary in freshwater to influence ecosystem functioning

Andrew J Tanentzap, Amelia Fitch, Chloe Orland, Erik JS Emilson, Kurt M Yakimovich, Helena Osterholz, Thorsten Dittmar

Corresponding author: Andrew J Tanentzap
Email: ajt@cam.ac.uk

This PDF file includes:

Figs. S1 to S6
Tables S1 to S3
References for SI reference citations

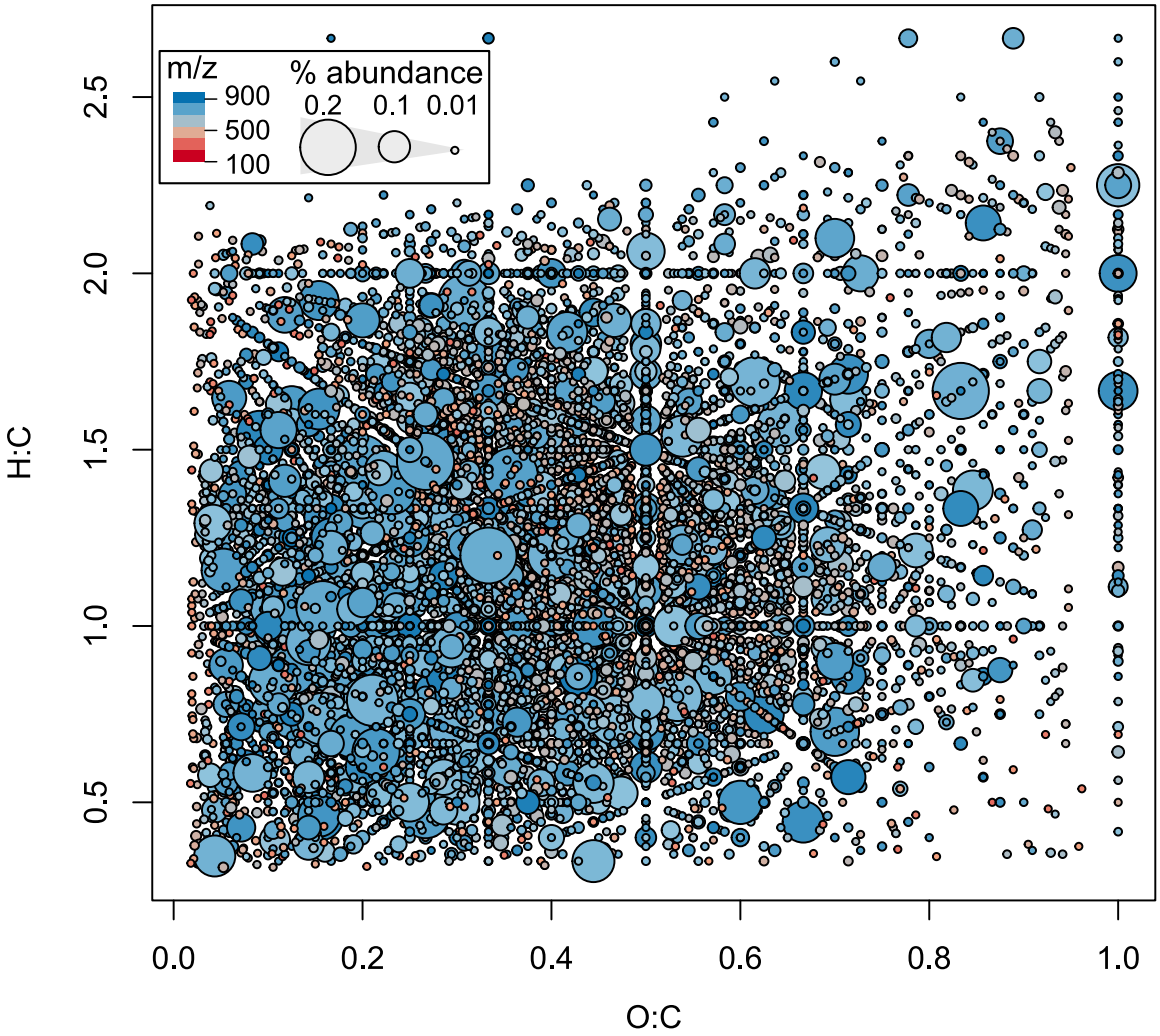

Fig. S1. Individual molecules varied in their chemical properties. Van-Krevelen plot shows 12,900 molecular formulas coloured according to their mass (m/z ratio, Da) and sized according to their percent abundance summed across 25 sediment mesocosms.

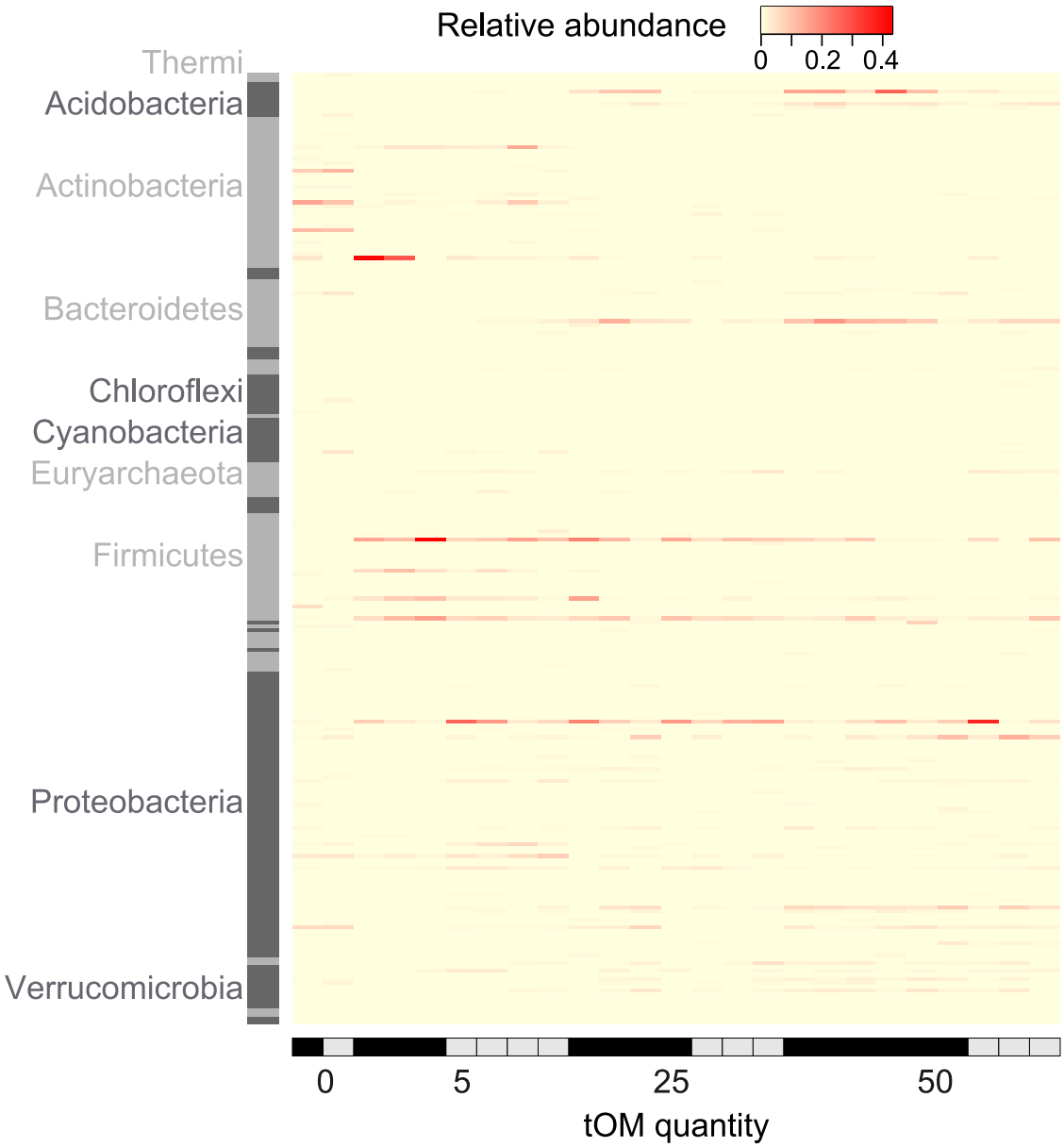

Fig. S2. Microbial community profile across mesocosms in the dark and light lake. Cells are relative abundances of 94 taxonomic families (rows) in each of 25 mesocosms (columns). Most OTUs (2,651 of 3,613) could be classified to the family-level. Row labels denote major microbial phyla corresponding with groups of families. Column labels denote quantity of terrestrial organic matter (tOM) added into the dark and light lake, respectively represented by dark and light shading.

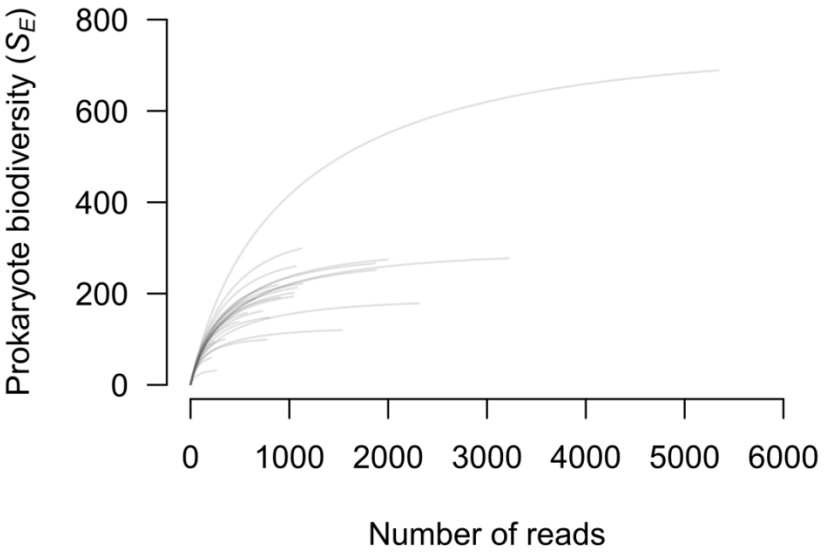

Fig. S3. Individual-based rarefaction curves show that we achieved representative sampling. Curves were calculated for operational taxonomic units (OTUs) in each sample with random sampling.

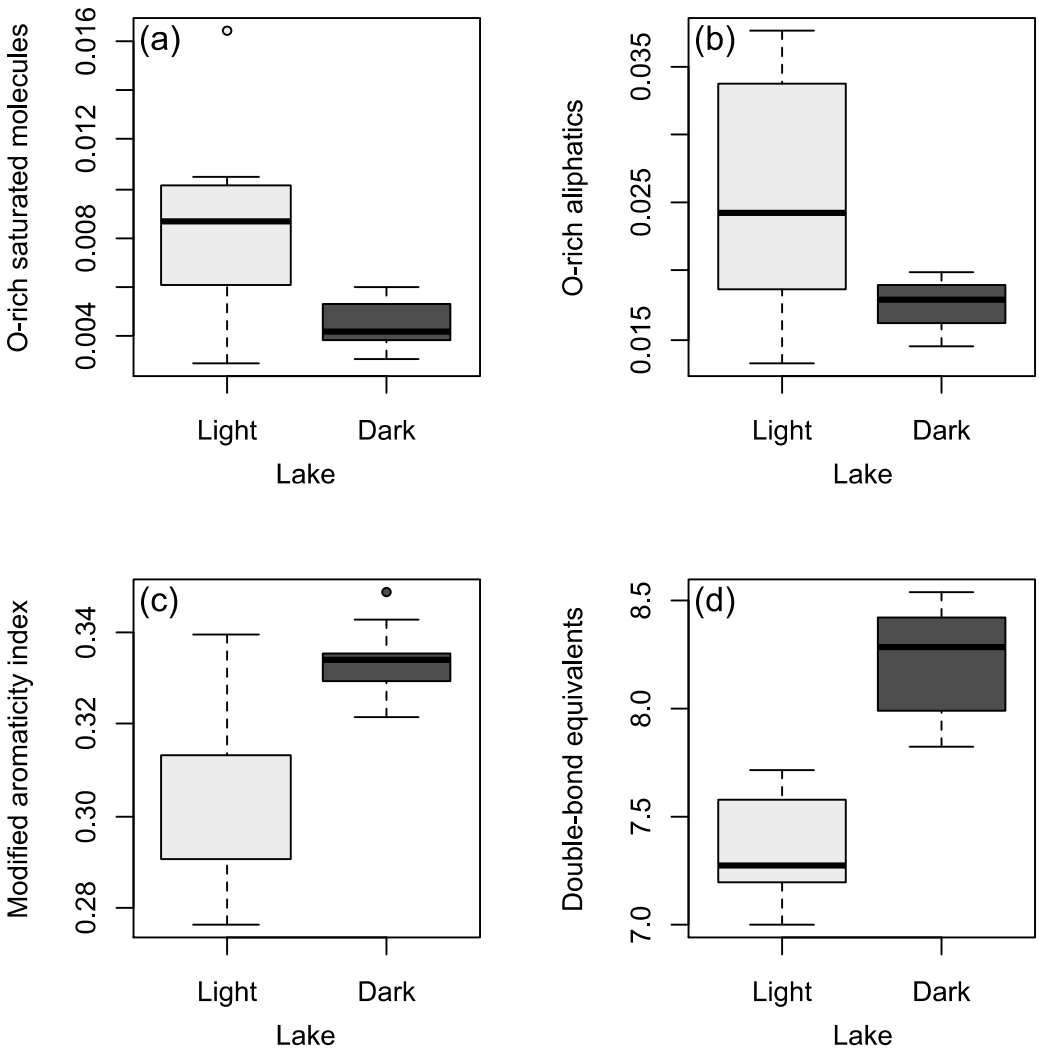

Fig. S4. Photo-oxidation indicators differed between pore water in mesocosms of the light and dark coloured lake. The light-coloured lake had a greater relative abundance of **(a)** oxygen-rich saturated molecules (two sample *t*-test: $t_{23} = 3.98$, $p < 0.001$) and **(b)** oxygen-rich unsaturated aliphatic molecules (Welch two sample: $t_{12.47} = 2.92$, $p = 0.013$). By contrast, mesocosms in the dark lake had a larger modified aromaticity index ($t_{13.66} = 4.70$, $p < 0.001$) and more double-bond equivalents ($t_{23} = 8.87$, $p < 0.001$) in mixtures of MFs, where means were weighted by the relative abundances of each molecule in each sample (i.e. community-weighted mean). Molecular formulas were assigned to compound groups after ref. 1, wherein saturated compounds had H/C>2, 0.5<O/C<0.9 and oxygen-rich aliphatics had 1.5<H/C<2, 0.5<O/C<0.9, N=0.

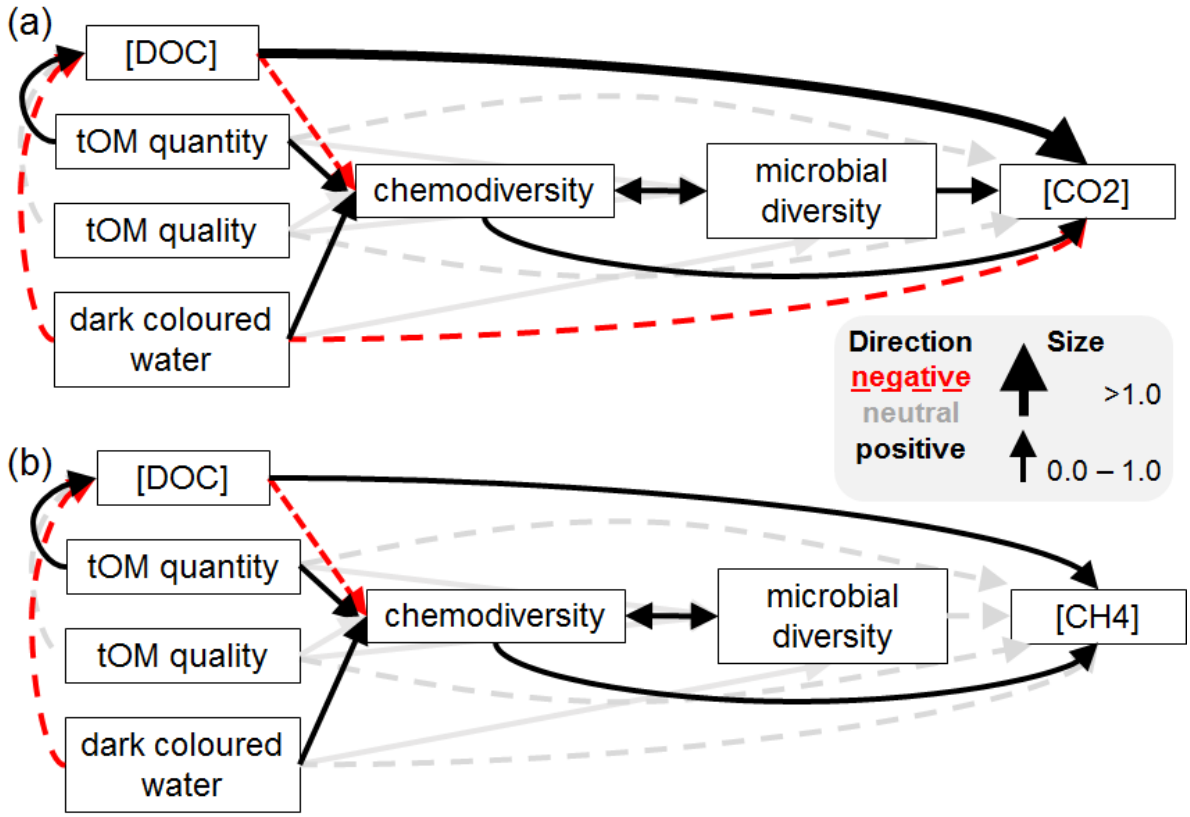

Fig. S5. Visualization of our statistical analyses as a path analysis. We refit the models described in the main text using the *piecewiseSEM* R package (ref. 2) for concentrations of **(a)** CO_2 and **(b)** CH_4. Arrows point at modeled responses, with mean effects of one variable on another proportional to their standardized effect size (see legend). As path analysis cannot include feedbacks between variables, we separately fitted the models with either chemodiversity or microbial diversity, and averaged effects across the two alternative models predicting each greenhouse gas concentrations for plotting. There were no missing linkages in any of the models and the results were indistinguishable from those in Table 1.

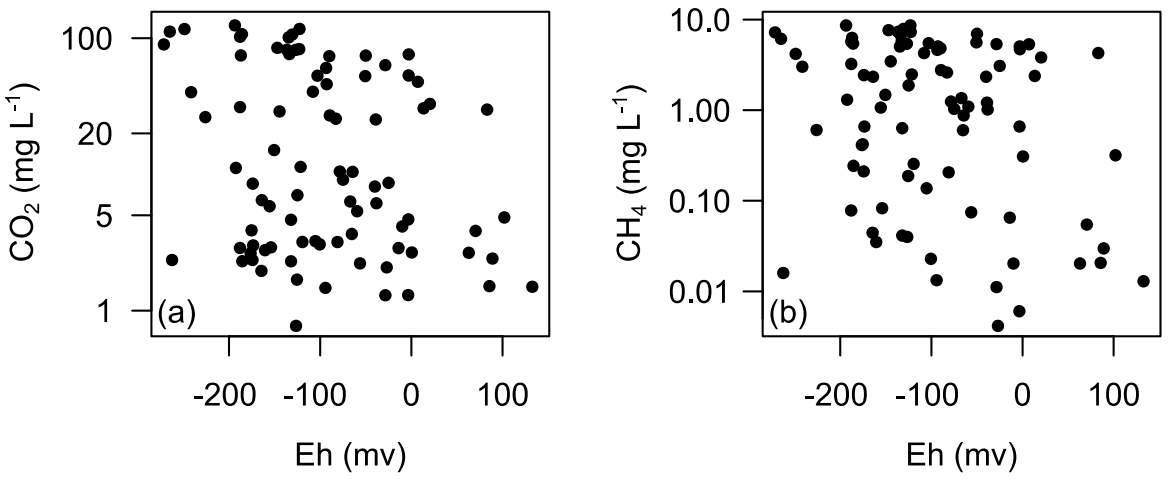

Fig. S6. Greenhouse gas concentrations were not higher in more anoxic sediments. We installed half-cell platinum electrodes designed to record *in situ* reduction-oxidation potential (Eh) in 12 mesocosms per lake during May 2016. Electrodes were built after ref. 3 from pure platinum wire encased in a 1 mL pipette tip filled with marine epoxy and 2.5 cm of exposed wire and were anchored with a small rock at the sediment bottom. Probe accuracy was tested as described by ref. 4 using a ZoBell's solution (Hach, London, Canada), and all probes had consistent error values relative to each other (standard deviation = 0.6 mV). We then recorded Eh opportunistically on 4 occasions over 2 months on dates coinciding with pore water sampling. Measurements were taken by placing a silver:silver chloride (with saturated KCl) reference probe into the overlying lake water and attaching it and the *in situ* platinum probe to a multimeter. Readings were allowed to stabilize over a 5-minute period. Eh was calculated by adding the measured electrical potential in millivolts to the electrical potential of the reference electrode given the overlying water temperature calculated after ref. 4. We then fitted a linear model to predict pore water concentrations of **(a)** CO_2 and **(b)** CH_4 given Eh, lake identity, dissolved organic carbon concentration, and sampling date, and we accounted for repeated measurements of the same mesocosm separately within each lake. These models showed no effect of Eh on either CO_2 ($t_{59.11} = 0.109$, $p = 0.915$) or CH_4 ($t_{69.44} = 0.632$, $p = 0.529$), as estimated with the *lmerTest* R package. There was also no difference in Eh between lakes when we paired identical mesocosms in our block design across each date (paired *t*-test: $t_{37} = 1.61$, $p = 0.112$).

Table S1 – Linear models for five responses describing mixtures of molecular formulas and two responses describing microbial communities. Responses included different measures of functional diversity (FD). For decomposition, FD was calculated for high-level categories (n = 4) and all genes associated with these (n = 726). Values in cells are mean estimates ± standard error. All predictors were scaled to a mean of zero and standard deviation of one so that effects are directly comparable. Bolded values are statistically significant at ***$p < 0.001$, **$p < 0.01$ *$p < 0.05$. Degrees of freedom = 20 and 21 for chemical and biological metrics, respectively.

	Model terms					
Response	Intercept	Dark lake	DOC	tOM quantity	tOM quality	R^2
Chemodiversity	**7.83 (0.03)*** **	**0.18 (0.04)*** **	**-0.10 (0.02)*** **	**0.06 (0.02)***	0.01 (0.03)	0.73
FD (MF size)	**17.1 (0.08)*** **	**0.45 (0.12)** **	-0.001 (0.07)	0.11 (0.07)	-0.16 (0.08)	0.43
FD (bioavailability)	**1.26 (0.01)*** **	**-0.07 (0.01)*** **	-0.01 (0.01)	-0.01 (0.01)	**-0.01 (0.01)***	0.81
FD (energetic rewards)	**0.47 (0.01)*** **	**-0.08 (0.02)*** **	**0.03 (0.01)*** **	-0.01 (0.01)	**-0.02 (0.01)***	0.76
FD (aromaticity)	**0.30 (0.003)*** **	**0.04 (0.004)*** **	**0.01 (0.002)*** **	0.002 (0.002)	0.01 (0.002)	0.84
Biodiversity	**4.92 (0.17)*** **	0.40 (0.24)	n/a	-0.04 (0.13)	0.09 (0.18)	0.01
FD (decomposition categories)	**0.82 (0.004)*** **	0.01 (0.01)	n/a	**0.01 (0.003)** **	0.004 (0.01)	0.39
FD (decomposition genes)	**5.14 (0.01)**	-0.03 (0.01)	n/a	**0.02 (0.01)***	-0.01 (0.01)	0.22

Table S2 – Chemodiversity and biodiversity were associated with each other. Cells are mean estimates ± standard error for terms in linear models separately predicting the Shannon-Wiener index for each type of diversity measure. All responses and predictors were scaled to a mean of zero and standard deviation of one so that effects are directly comparable between diversity indices. Bolded values are statistically significant at ***$p < 0.001$; **$p < 0.01$, degrees of freedom = 22.

	Response	
Model term	Chemodiversity	Biodiversity
Intercept	**-0.63 (0.18)****	0.19 (0.31)
Dark lake	**1.21 (0.25)*****	-0.37 (0.50)
Chemodiversity	n/a	**0.73 (0.28)****
Biodiversity	**0.38 (0.13)****	n/a
R^2	0.65	0.30

Table S3 – Linear models predicting log-transformed CO_2 and CH_4 with each of four functional diversity metrics for DOM. Values in cells are mean estimates ± standard error. All predictors were scaled to a mean of zero and standard deviation of one so that effects are directly comparable. Bolded values are statistically significant at ***$p < 0.001$, **$p < 0.01$ *$p < 0.05$. Degrees of freedom = 19.

		Model terms					
FD predictor	Intercept	FD metric	Dark lake	DOC	tOM quantity	tOM quality	R^2
			CO_2				
MF size	**2.59 (0.24)*** **	-0.25 (0.20)	0.16 (0.39)	**1.16 (0.17)*** **	0.07 (0.17)	-0.20 (0.22)	0.77
Bioavailability	**1.93 (0.32)*** **	**0.87 (0.29)** **	0.98 (0.55)	**1.24 (0.15)*** **	0.13 (0.15)	0.18 (0.19)	0.83
Energetic rewards	**2.42 (0.27)*** **	0.51 (0.30)	0.08 (0.43)	**0.92 (0.22)*** **	0.08 (0.16)	0.08 (0.22)	0.79
Aromaticity	**1.88 (0.36)*** **	**-0.91 (0.33)***	1.09 (0.62)	**1.66 (0.24)*** **	0.09 (0.15)	0.12 (0.19)	0.82
			CH_4				
MF size	0.25 (0.41)	-0.42 (0.35)	0.44 (0.68)	**1.27 (0.30)*** **	0.29 (0.30)	-0.46 (0.38)	0.55
Bioavailability	**-1.15 (0.51)***	**1.78 (0.46)** **	**2.87 (0.87)** **	**1.45 (0.24)*** **	0.44 (0.23)	0.27 (0.31)	0.73
Energetic rewards	-0.19 (0.46)	**1.08 (0.50)***	1.08 (0.72)	0.75 (0.37)	0.34 (0.28)	0.09 (0.37)	0.62
Aromaticity	**-1.24 (0.58)***	**-1.85 (0.54)** **	**3.08 (1.01)** **	**2.28 (0.38)*** **	0.36 (0.24)	0.15 (0.31)	0.71

References

1. Seidel M, et al. (2014) Biogeochemistry of dissolved organic matter in an anoxic intertidal creek bank. *Geochim Cosmochim Acta* 140:418–434.
2. Lefcheck JS (2016) piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods Ecol Evol* 7:573–579.
3. Swerhone GDW, Lawrence JR, Richards JG, Hendry MJ (1999) Construction and testing of a durable platinum wire Eh electrode for *in situ* redox measurements in the subsurface. *Groundwater Monitoring & Remediation* 19:132–136.
4. Nordstrom DK, Wilde, FD. 2005. Reduction oxidation potential (electrode method), v1.2 in *U.S. Geological Survey Techniques of Water-Resources Investigations, book 9*. (U.S. Geological Survey), pp.1-20.